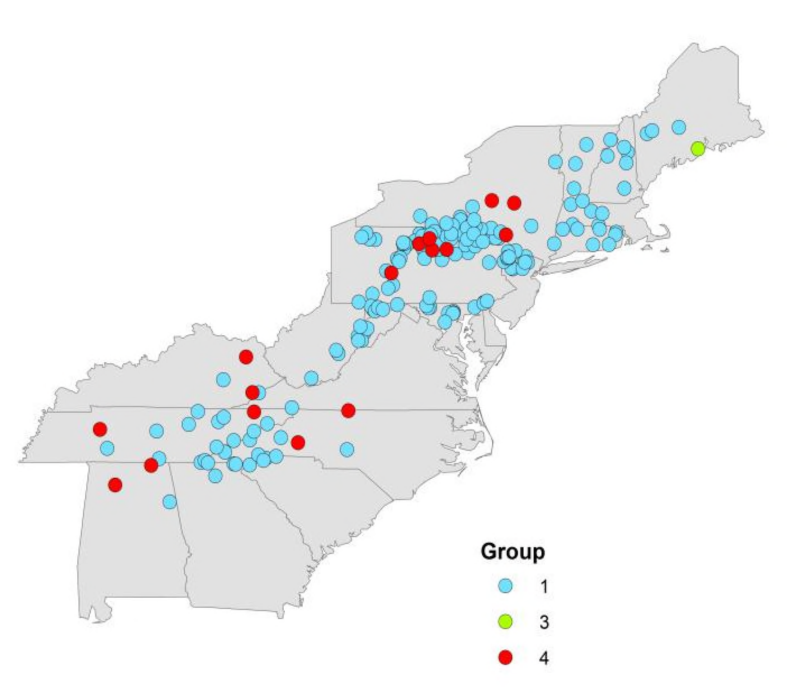
Stream Classification in Support of Regional Monitoring to Detect Climate Change Effects

National Water Quality Monitoring Conference

Session E4
April 29, 2014

Jonathan Witt & Britta Bierwagen,
ORD/USEPA
Jen Stamp & Anna Hamilton,
Tetra Tech, Inc.



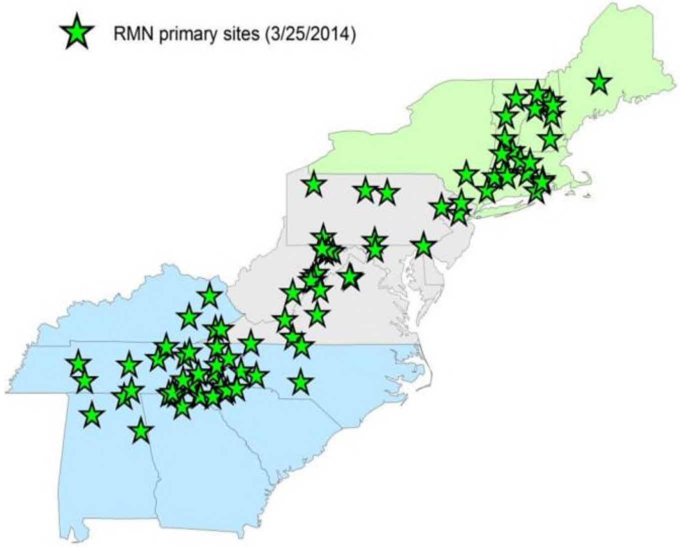
Regional Monitoring Networks

- Stream impairment is often determined by comparing macroinvertebrate communities between reference and non-reference streams

- With changing climate reference communities may shift, which influences baseline conditions used in assessments

Regional Monitoring Networks

- Working with states and EPA Regional offices in Northeast, Mid-Atlantic, Southeast to develop regional monitoring networks (RMNs)



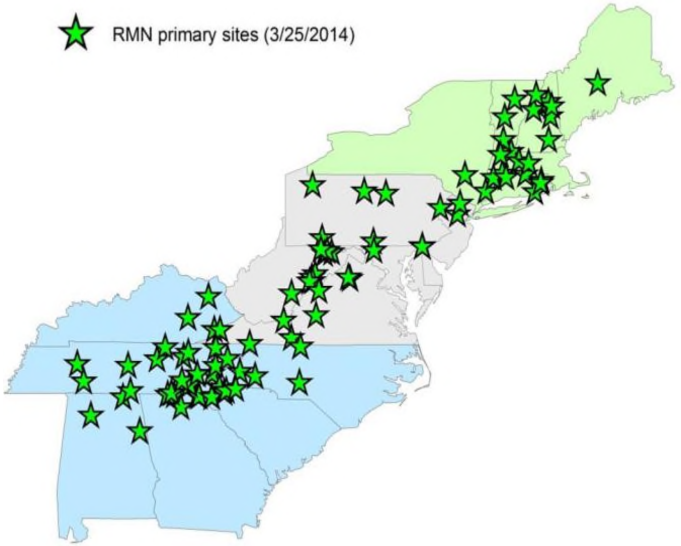
- Purpose is to detect climate-related changes and establish reference condition baselines

- Focus on quantifying stream biota, thermal and hydrologic regimes in minimally disturbed freshwater wadeable streams over time

Objectives

Create a biological classification

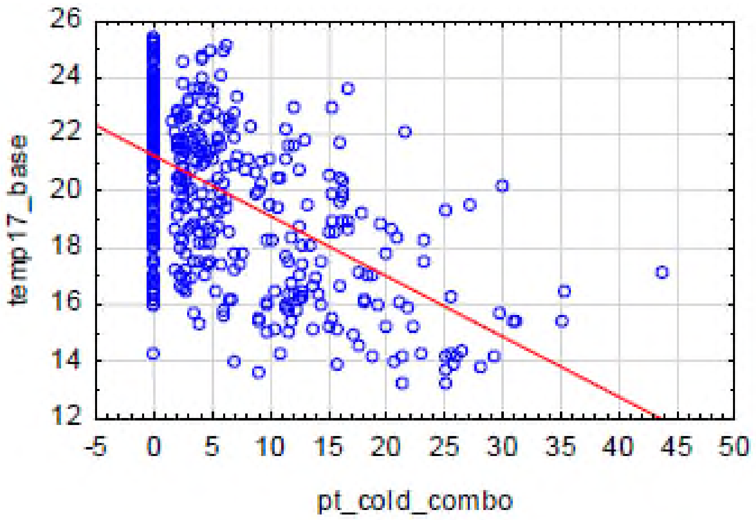
- Allow integration of sites by community types across regions
  - o Climate change occurs across a large scale
  - o Program resources are limited
- Inform site choice for current and future monitoring
  - o Where community changes are expected
  - o To allow validation and testing of climate-related hypotheses

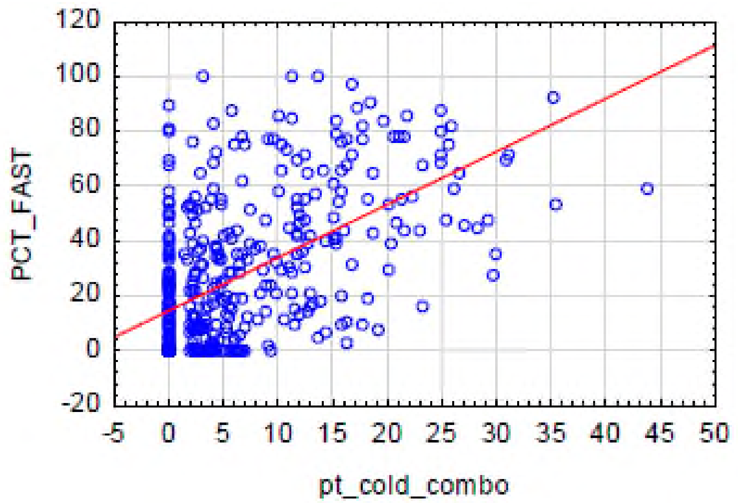


Objectives

Determine which environmental variables within and across regions are the best predictors of community composition

- Broad scale habitat variables, e.g. temperature, precipitation and flow
- Use variables to predict community groups





Methodology

DATA SOURCE

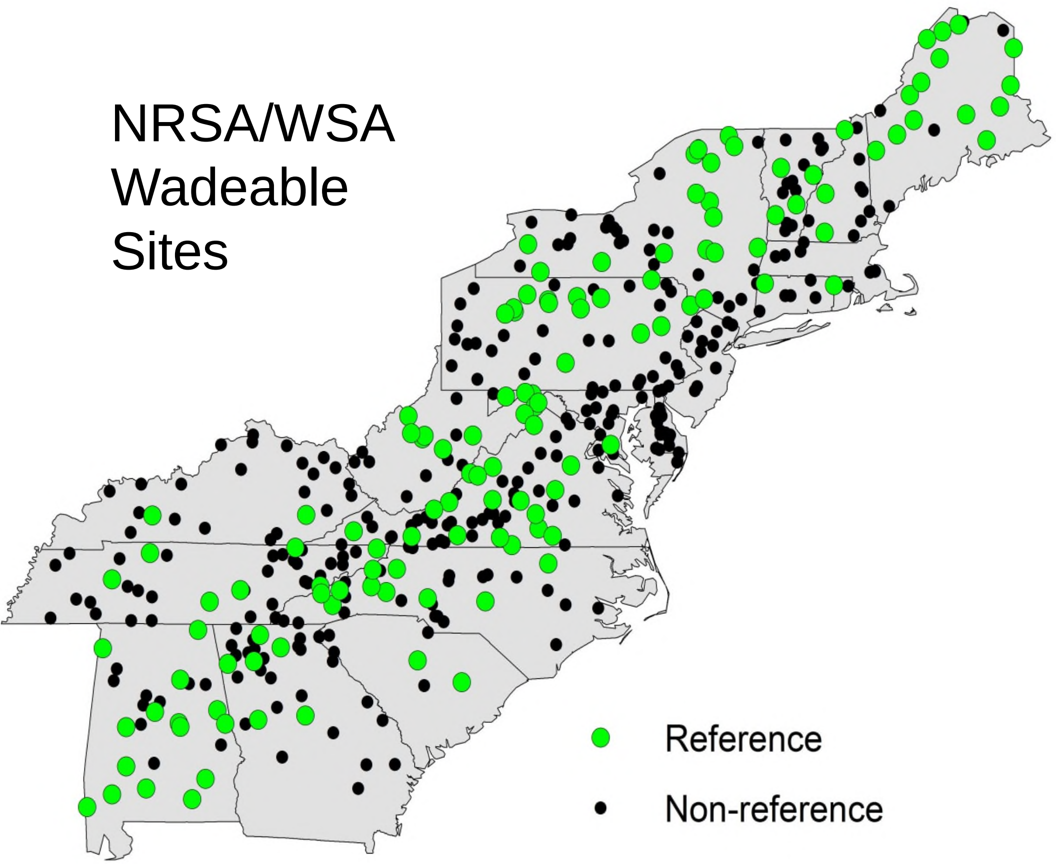
- Wadeable stream data from the 2010-2011 **National Aquatic Resource Surveys**
  - Standardized collection and processing methods
  - Available for entire study area
- Performed analyses on datasets comprised of 126 reference sites

ANALYSES

1. Cluster analysis on OTU presence-absence
   - to identify community groups
2. Multinomial GAMs
   - predict community group with environmental variables
3. Multimodel Inference
   - increases robustness

Methodology
Reference Criteria

NRSA/WSA
Wadeable
Sites



- % Urban < 5%
- % Row crops < 15%
- % Pasture < 25%
- Conductivity < 500
- pH > 5
- No dams within 1 km (upstream or downstream)
- No NPDES Major discharges, Superfund National Priorities List or Toxic Release Inventory (TRI) sites within 1 km upstream

Many of the reference sites are located in the highlands/mountains. Some areas are not represented in our dataset (e.g., PA limestone, NJ Pine Barrens)

Three Community Clusters

United States Environmental Protection Agency

G1: Small Steep Cold

- BIOLOGY: More EPTs, More Cold Water Taxa
- Generally streams < 100 km2, higher slope (>1.0%)
- Cooler temperatures (summer air temp < 17.5 C)
- Substrate - more % fast habitat (>40%)

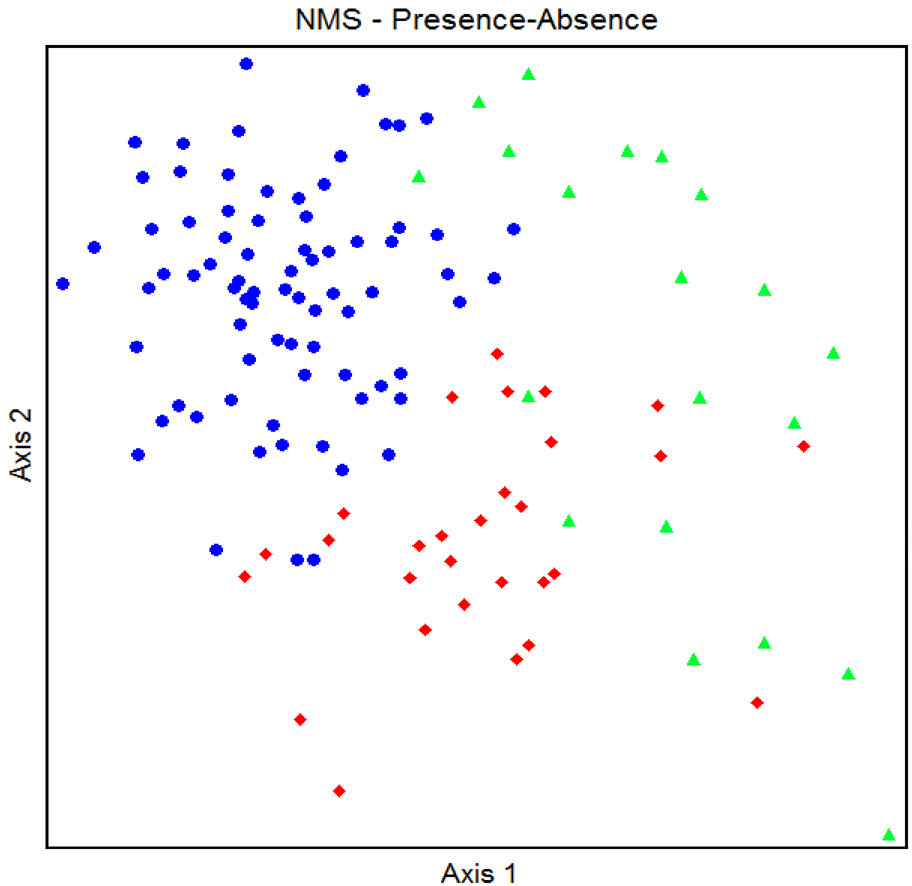
G3: Small Flat Warm

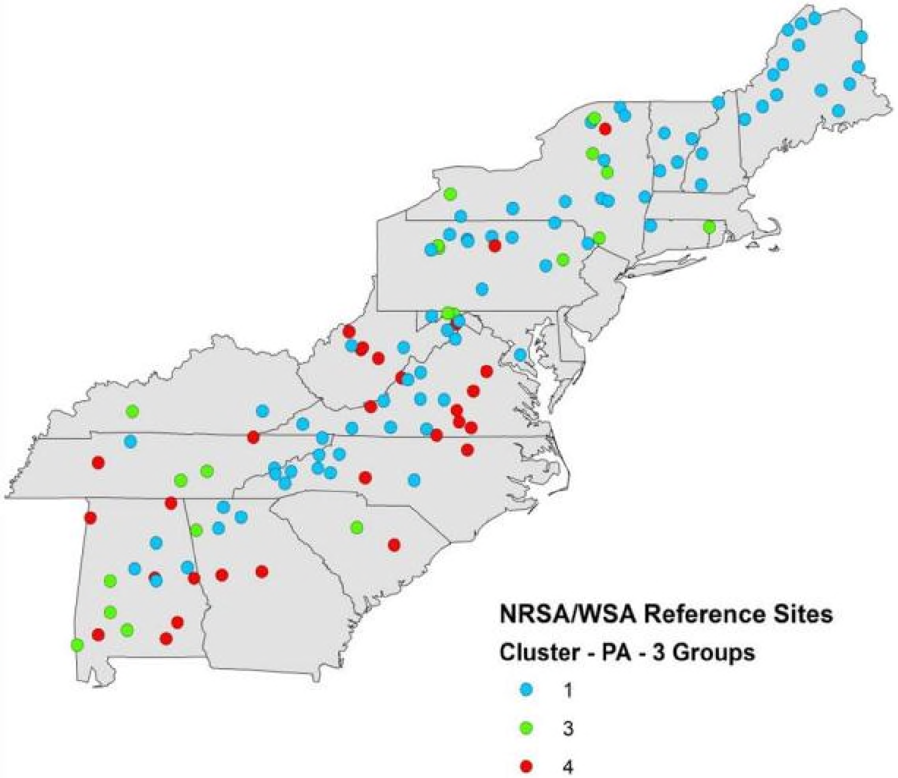
- BIOLOGY: Fewest EPTs, More Warm Water Taxa, More Chironomids
- Generally streams < 10 km2, lower slope (<1.0%)
- Substrate – more % fines (>5%), more % slow (>57%), and more % pool (>15%)

G4: Large Flat Warm

- BIOLOGY: Less EPTs, Most Warm Water Taxa
- larger drainage area, lower slope (<0.05)
- warmer temperatures (summer air temp >19 C)
- Substrate – more % sand (>5%), more % slow (>80%)

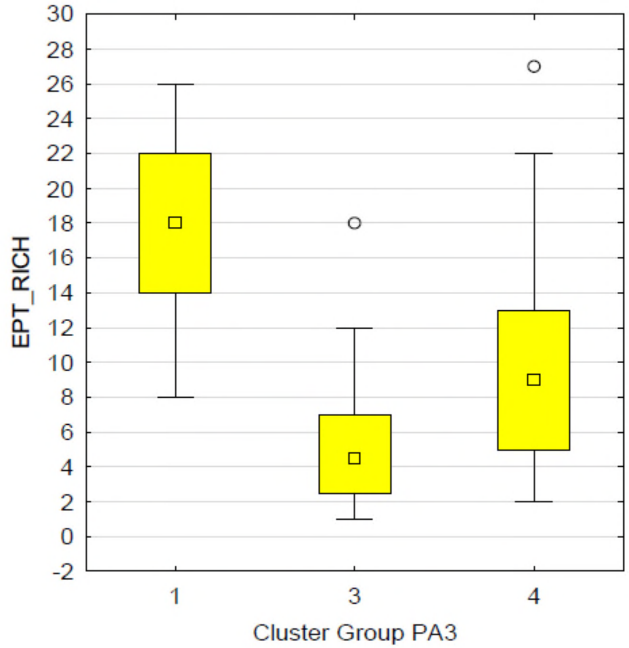
Three Community Clusters



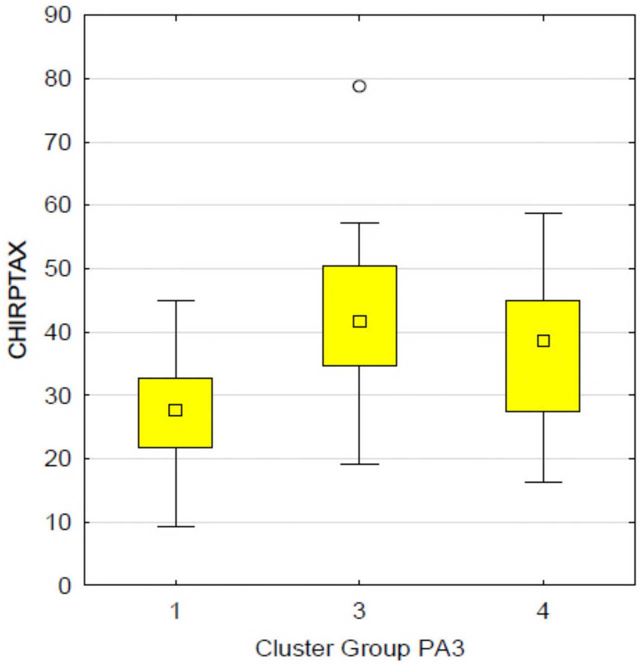


Three Community Clusters

EPT Richness

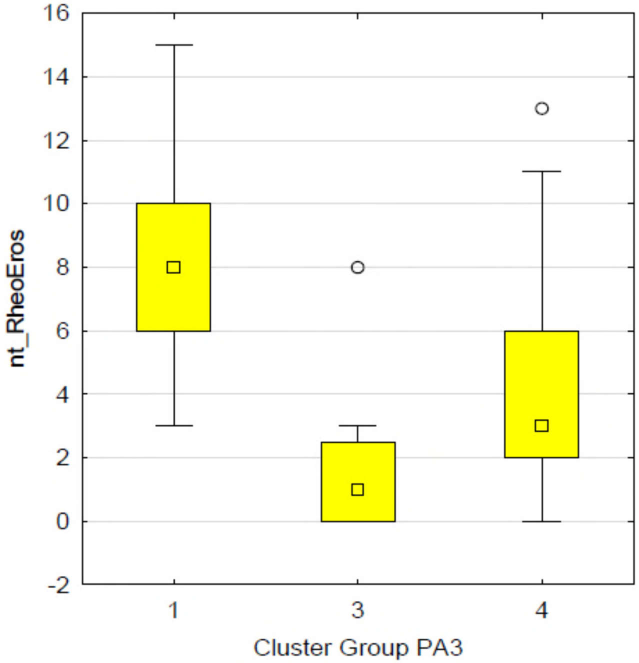


% Chironomid

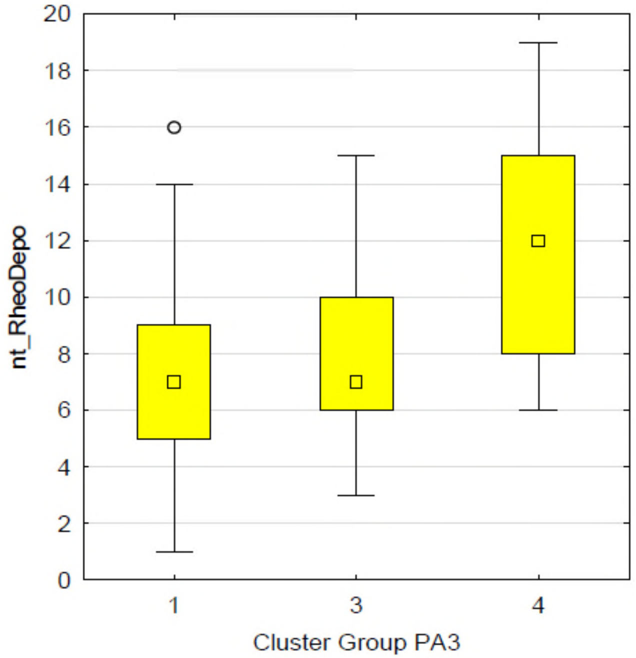


Three Community Clusters

Erosional Pref Richness

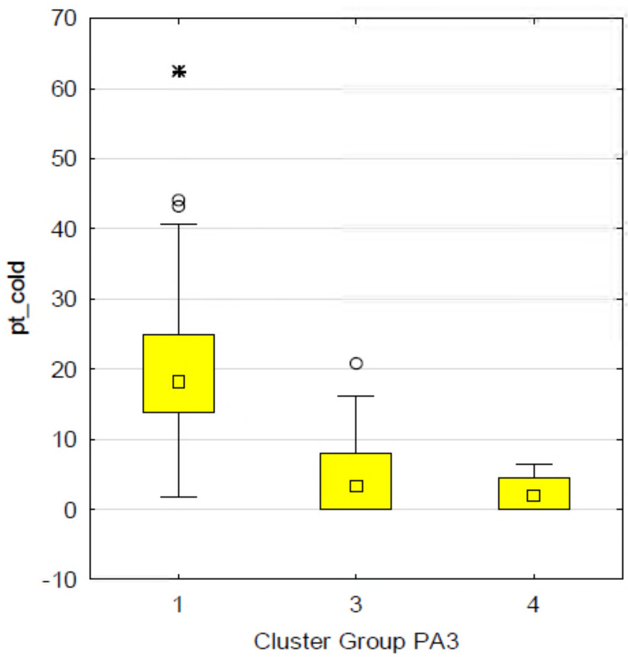


Depositional Pref Richness

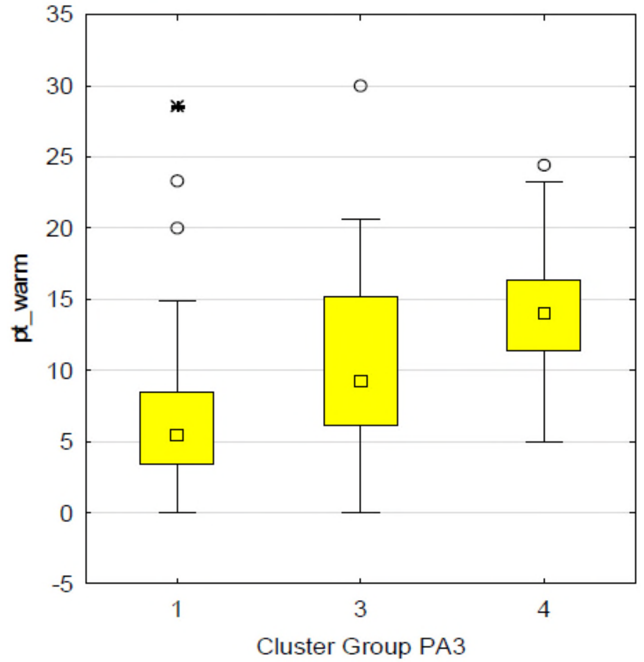


Three Community Clusters

% Cold Pref



% Warm Pref



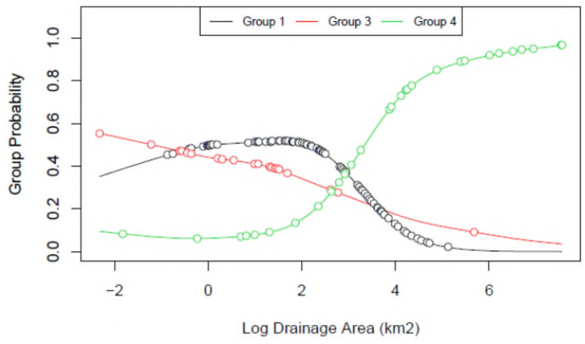
United States Environmental Protection Agency

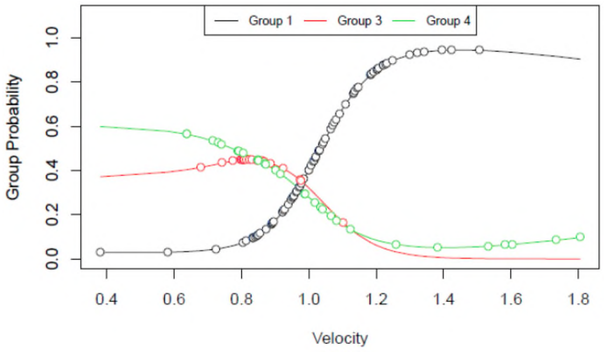
METHODOLOGY

GOAL: Predict probability of community class membership for catchments using a combination of environmental drivers.

Multinomial GAMs

- Categorical response variable (community group), similar to logistic regression but with more than two categories.
- Models the probability of being in a particular group
- GAMs allow for non-linear relationship, which is common in environmental data





Model Selection

- Outcome similar to discrimination and random forests
- Running models with different variables allows us to see which are most important for predicting class membership

METHODOLOGY

Used AICc and model averaging to predict class membership

In model averaging you don't select just one model, but a subset of models, where your response is a weighted average subset

Weighting still allows you to visually examine continuous patterns in the most important models, but also allows building of importance values

STEPS

1) Select variables for the models: which variables perform best

Catchment-level variables that can be extracted from NHDPlus and Climate Wizard* allow us to predict across regions

Drainage Area Elevation Baseflow Velocity Catchment Slope	Winter Temp Spring Temp Summer Temp Fall Temp
	Winter Precip Spring Precip Summer Precip Fall Precip

*1961-1990 Baseline Averages

Variable selection

The ensemble includes nine models with the following variables, listed in order of importance:

Velocity
Drainage Area
Summer Temp
Winter Temp
Fall Precip
Winter Precip
Catchment Slope
Spring Precip

STEPS

2) Apply models to predict site class membership

COMID	G1	G3	G4	G_Dom
12107044	0.81	0.12	0.07	G1
12107046	0.73	0.16	0.11	G1
12107048	0.53	0.30	0.17	G1
12107050	0.77	0.14	0.09	G1

- Correct Predictions: 84% (using highest probability group)
  - Under-represents G3 sites

- The model output does not provide any information on which taxa comprise the predicted class/assemblage directly. Rather we infer that from the group itself.

STEPS

3) Test performance on a preliminary RMN dataset

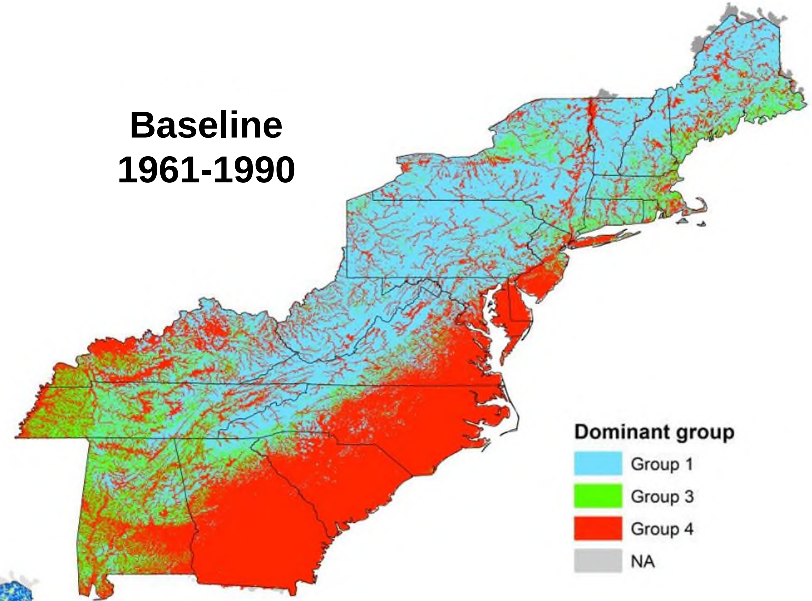
- Do observed communities there align with what the model predicts?
- In progress

4) Apply model to NHDPlus catchments in the Eastern US

- Climate change vulnerability assessment – how much does the probability of membership in G1 change if future projected temperature and precipitation variables are plugged into the model

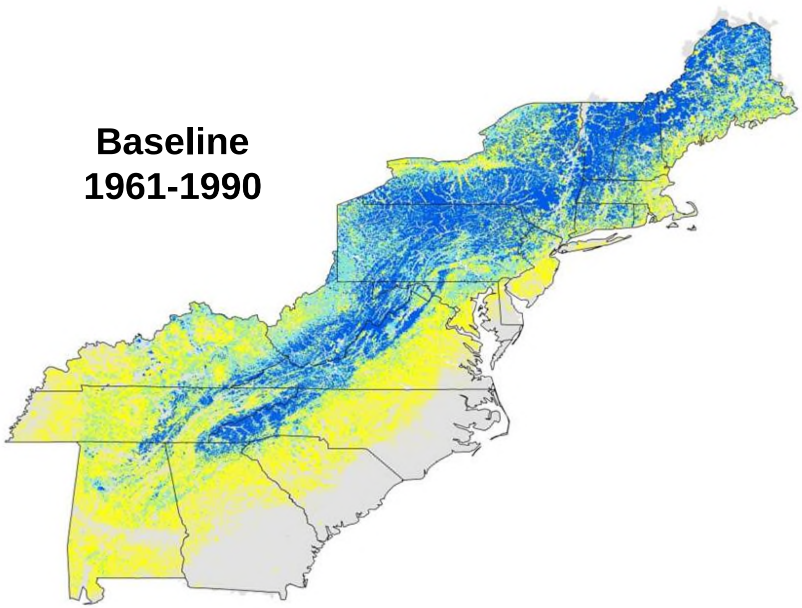
United States Environmental Protection Agency

Baseline 1961-1990



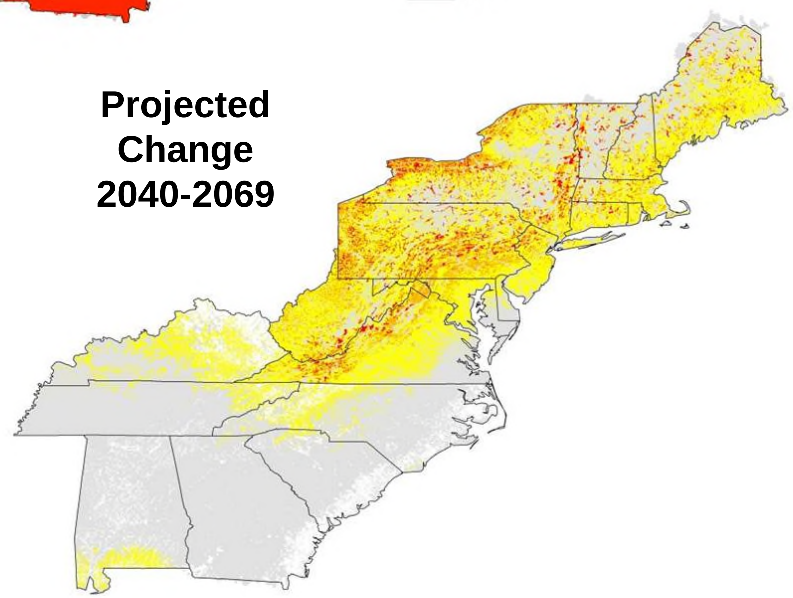
Dominant group
Group 1
Group 3
Group 4
NA

Baseline 1961-1990



Probability of membership in Group 1
G1
0 - 0.24
0.25 - 0.49
0.50 - 0.74
0.75 - 1.00

Projected Change 2040-2069



Change in Group 1 membership (mid-century minus baseline)
G1_diff
-0.22 to -0.15
-0.14 to -0.10
-0.09 to -0.05
-0.04 to 0.05

Climate change vulnerability assessment

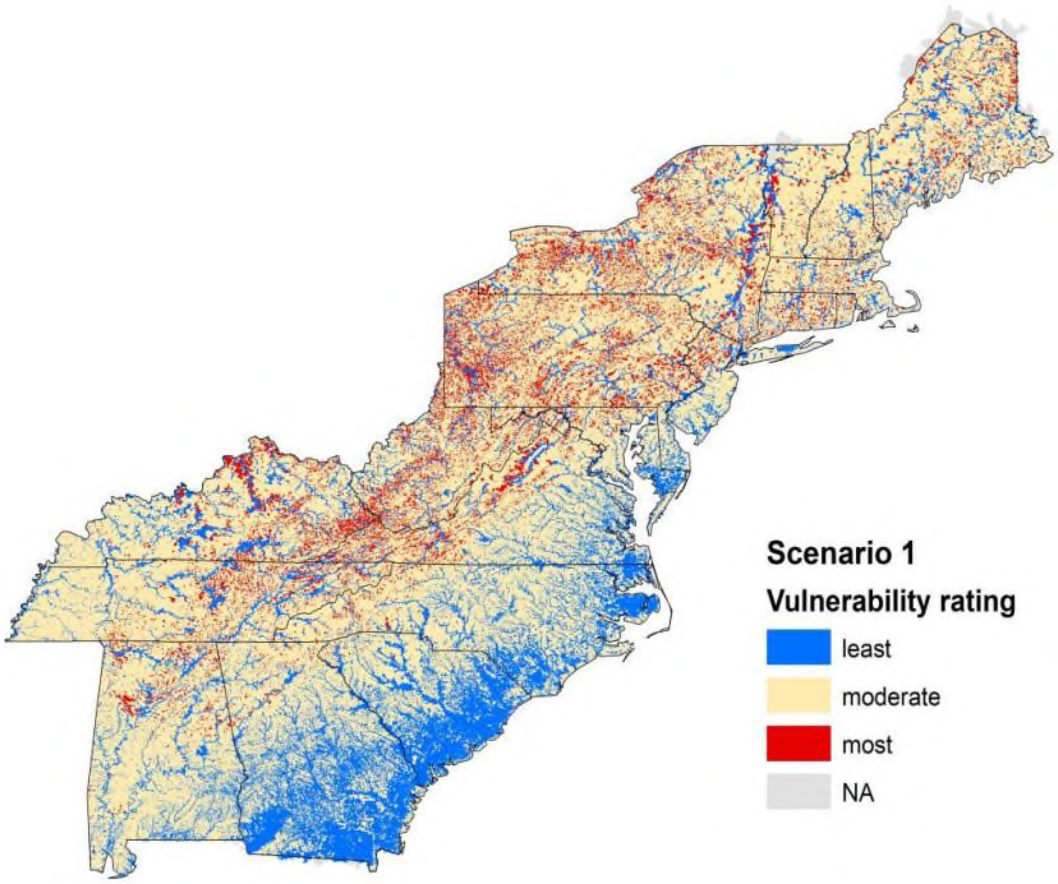
We rated the vulnerability of NHDPlus v1 local catchments to increasing summer temperatures based on the following:

Exposure

- ***Amount of projected change*** (mid-century minus baseline)
- ***Rate of change*** (climate velocity)
- ***Shading***
  - Riparian
  - Local catchment
  - Total watershed
- ***Baseflow***
- ***Urban*** land use (medium and high intensity)

Sensitivity

- ***Macroinvertebrate*** assemblage
  - Probability of membership in G1
  - Predicted change in G1 probability



Conclusions

- Biologically based classification allows us to aggregate data across the three RMNs, which span multiple ecoregions and other classification systems

- Assists with site selection for current and future monitoring efforts, and facilitates the testing of climate related hypotheses. For example,
  - Cold water taxa are expected to see range contractions with increasing water temperature
  - If sediment loads are expected to increase, we should expect taxa that require erosional environments (rheophiles) to decrease
  - Both of these taxa groups are found more frequently in G1 sites

Questions?

witt.jonathan@epa.gov
bierwagen.britta@epa.gov
jen.stamp@tetratech.com
anna.hamilton@tetratech.com

Multimodel Results

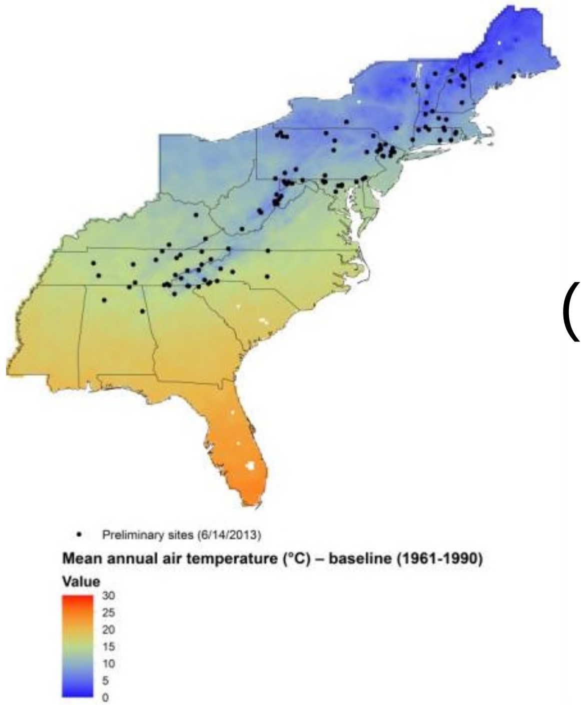
Model	ΔAICc	Evidence Ratio	Velocity	Drainage Area	Summer Temp	Winter Temp	Fall Precip	Winter Precip	Catchment Slope	Spring Precip
m1	0	1	0.26	0.26		0.26		0.26		
m2	0.34	1.19	0.22	0.22	0.22					
m3	1.20	1.83	0.14	0.14	0.14		0.14			
m4	1.42	2.04	0.13	0.13			0.13			
m5	1.74	2.39	0.11	0.11		0.11				0.11
m6	3.30	5.22	0.05		0.05		0.05		0.05	
m7	4.18	8.08	0.03		0.03				0.03	
m8	4.29	8.55	0.03	0.03						
m9	4.40	9.01	0.03		0.03			0.03	0.03	
Importance			**1.00**	**0.89**	**0.47**	**0.37**	**0.32**	**0.29**	**0.11**	**0.11**

Limited to models with evidence ratios < 10
Some redundant models were removed

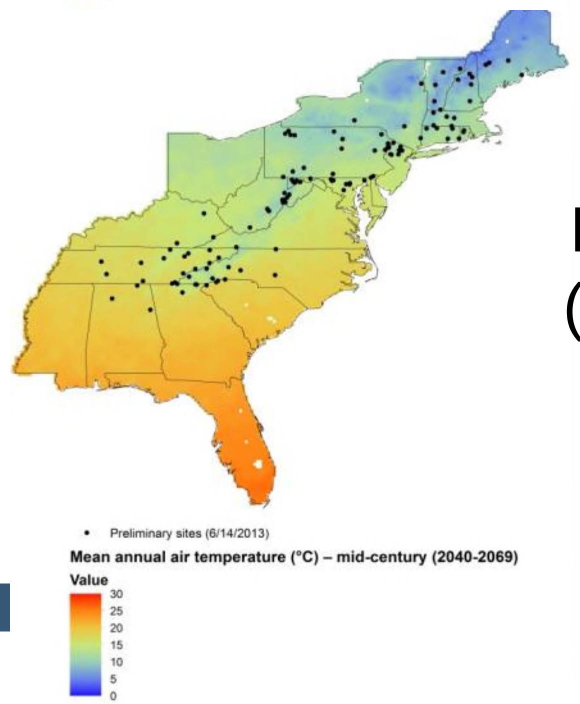
Correct Predictions: 84% (using highest probability group)
Under-represents G3 sites

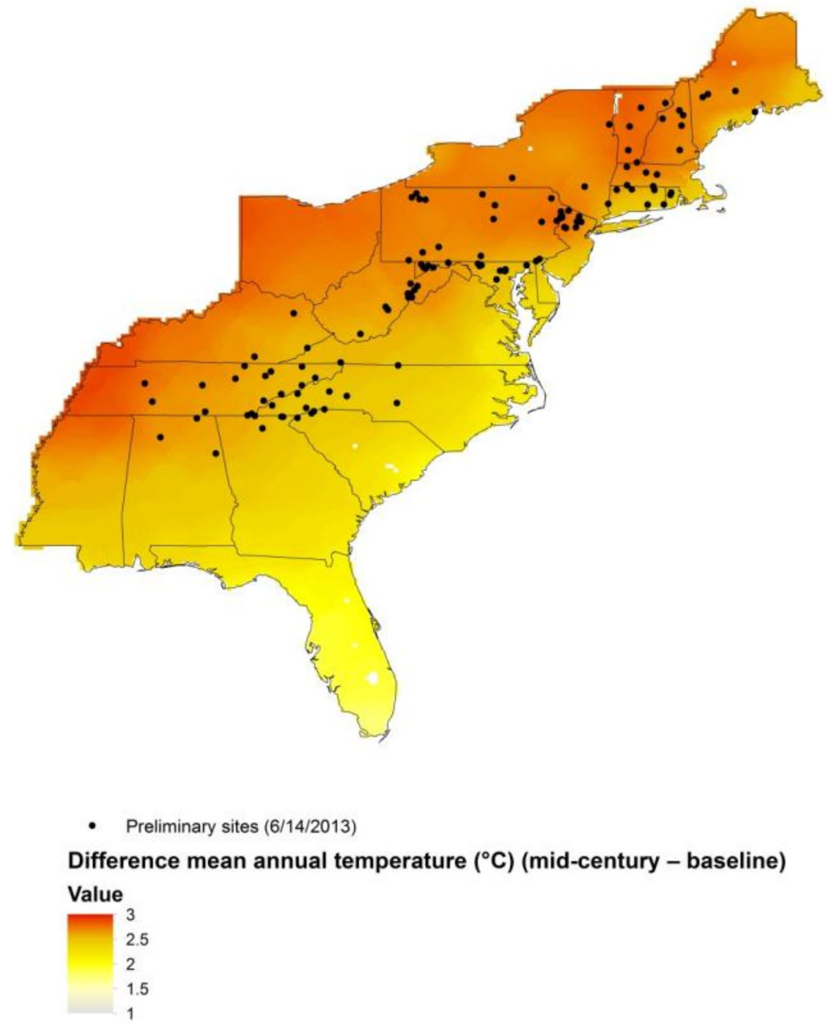
Air temperature – Mean Annual

Baseline
(1961-1990)

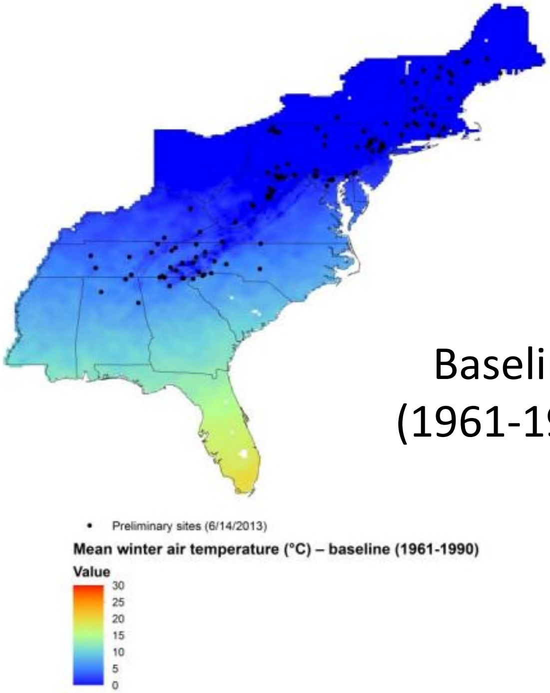


Mid-century
(2040-2069)

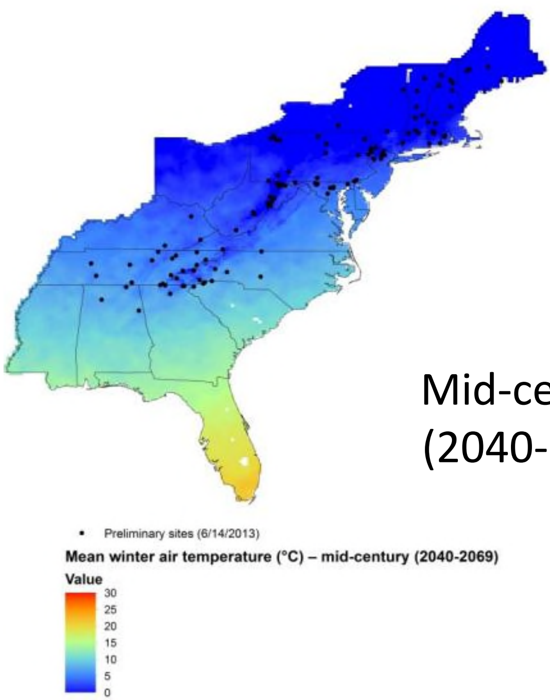




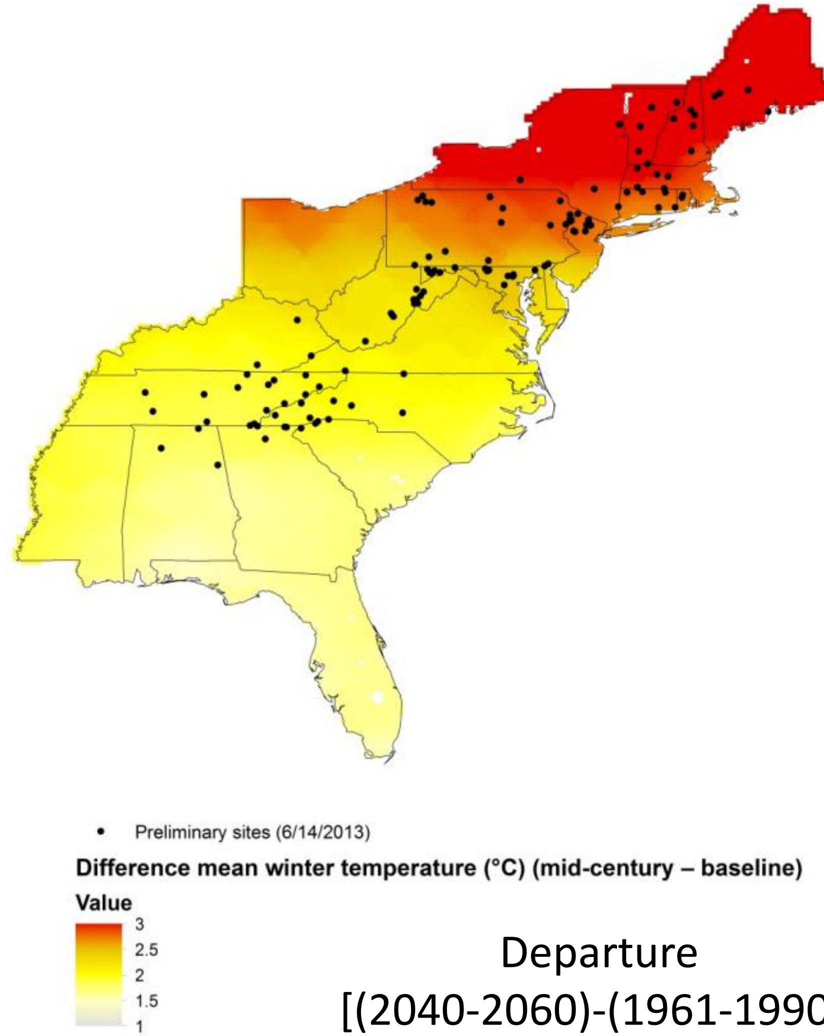
Air temperature – Mean winter (Dec-Feb)



Baseline (1961-1990)

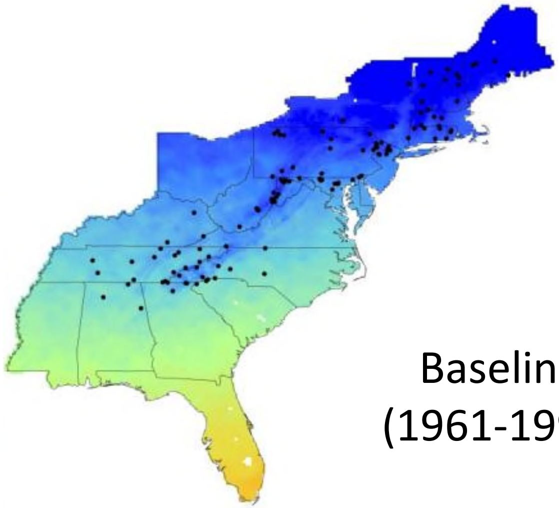


Mid-century (2040-2069)



Departure [(2040-2060)-(1961-1990)]

Air temperature – Mean spring (Mar-May)

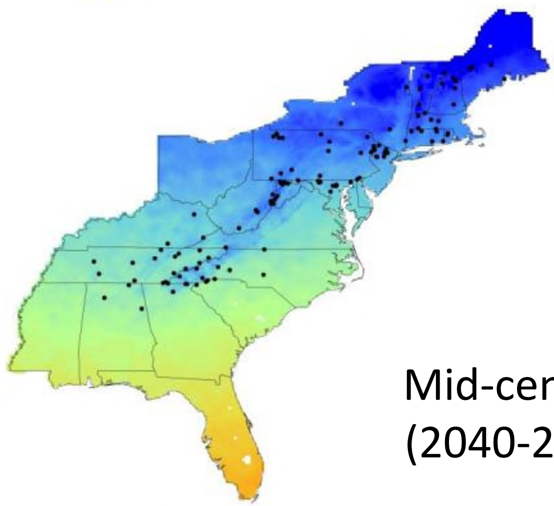


Baseline (1961-1990)

• Preliminary sites (6/14/2013)

Mean spring air temperature (°C) – baseline (1961-1990)

Value: 30, 25, 20, 15, 10, 5, 0

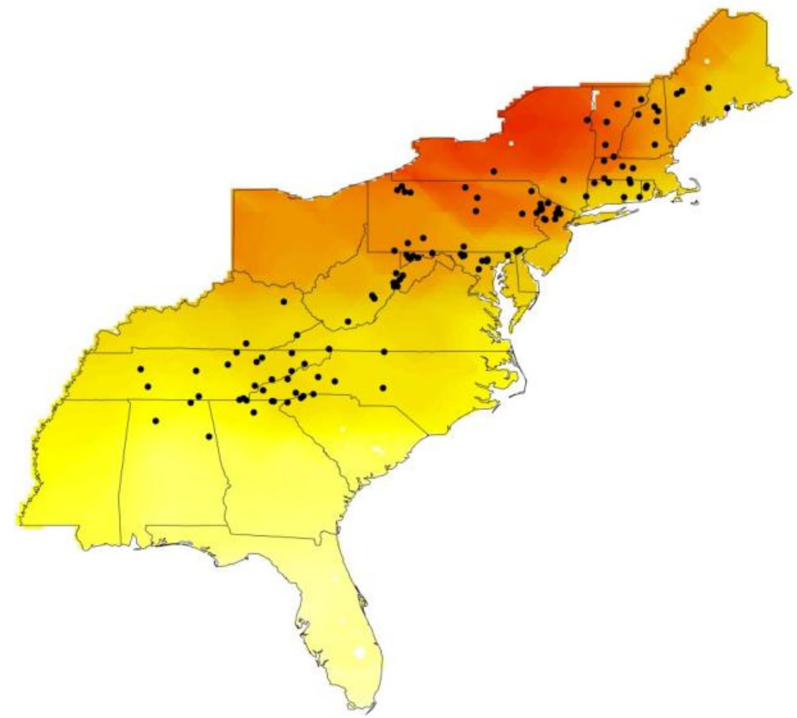


Mid-century (2040-2069)

• Preliminary sites (6/14/2013)

Mean spring air temperature (°C) – mid-century (2040-2069)

Value: 30, 25, 20, 15, 10, 5, 0



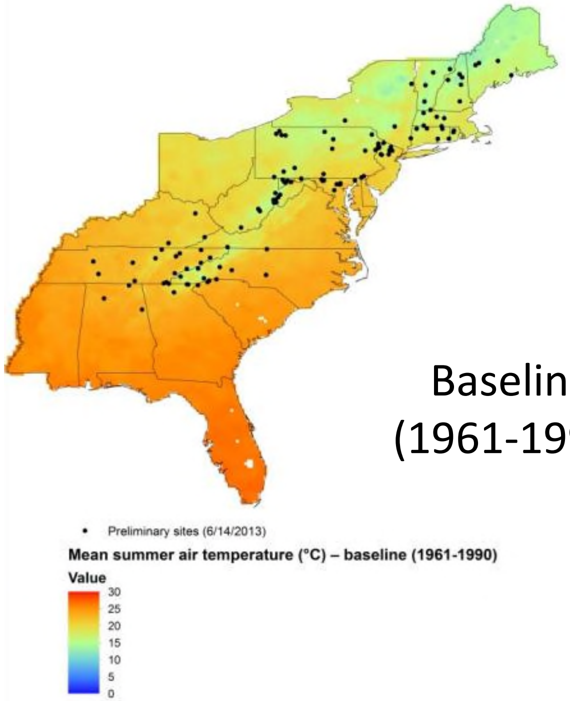
• Preliminary sites (6/14/2013)

Difference mean spring temperature (°C) (mid-century – baseline)

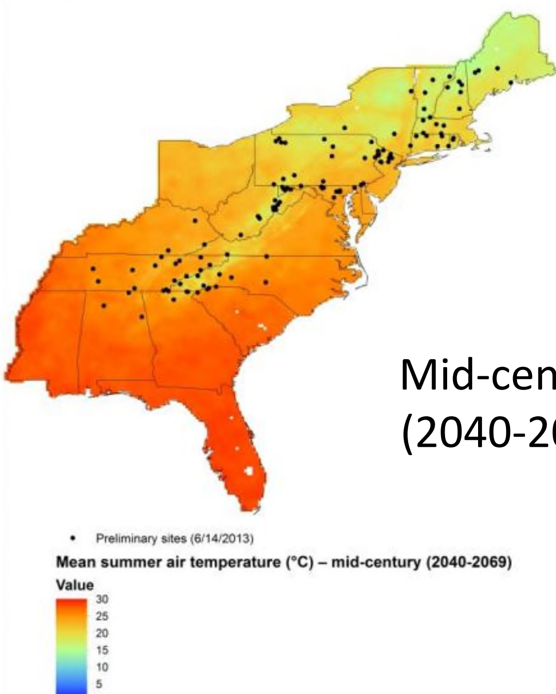
Value: 3, 2.5, 2, 1.5, 1

Departure [(2040-2060)-(1961-1990)]

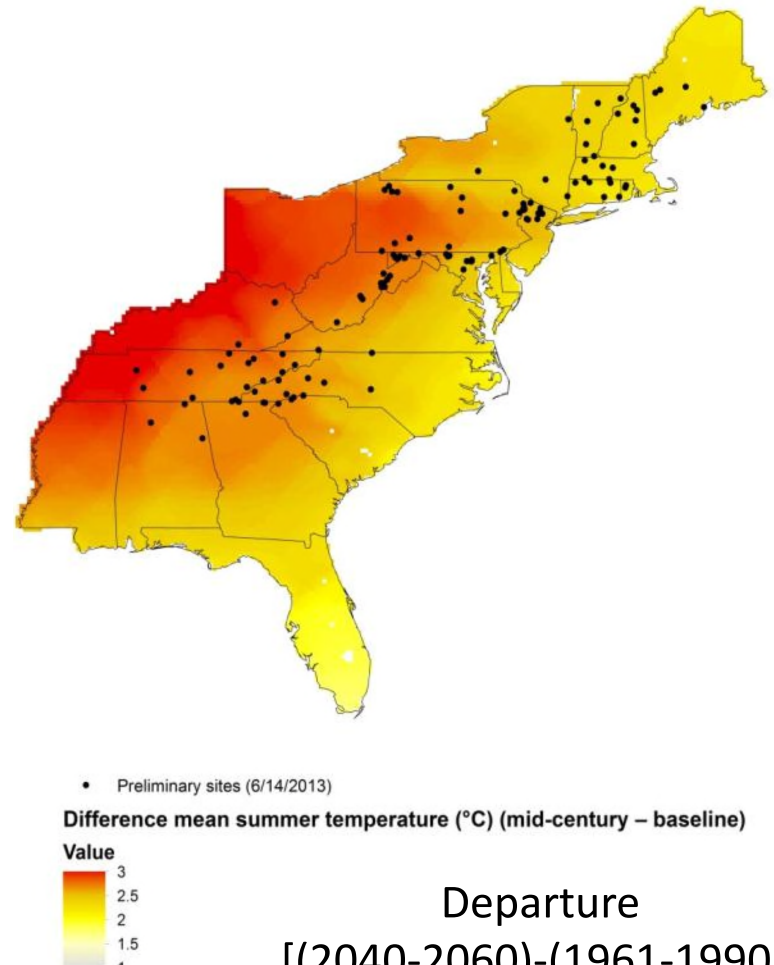
Air temperature – Mean summer (Jun-Aug)



Baseline
(1961-1990)

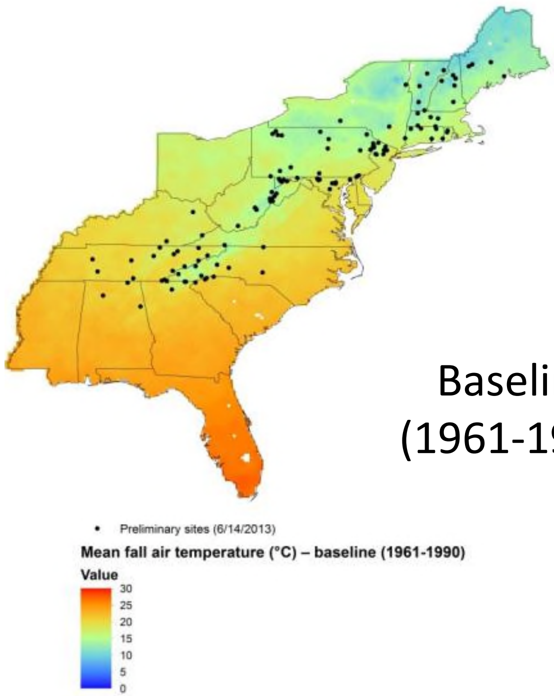


Mid-century
(2040-2069)

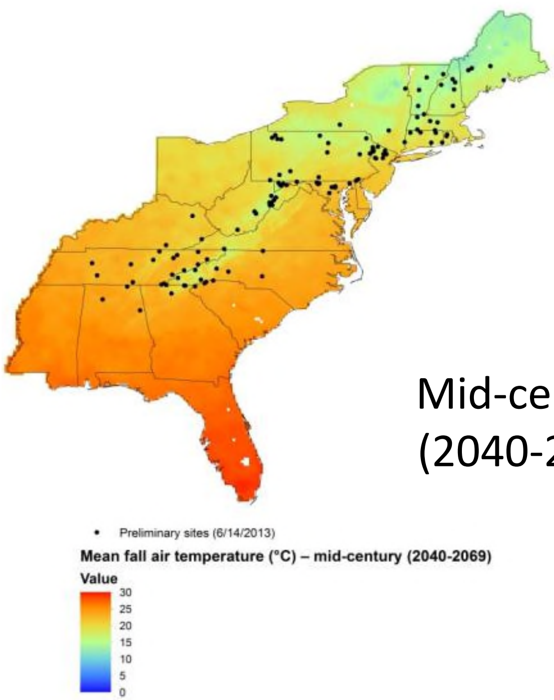


Departure
[(2040-2060)-(1961-1990)]

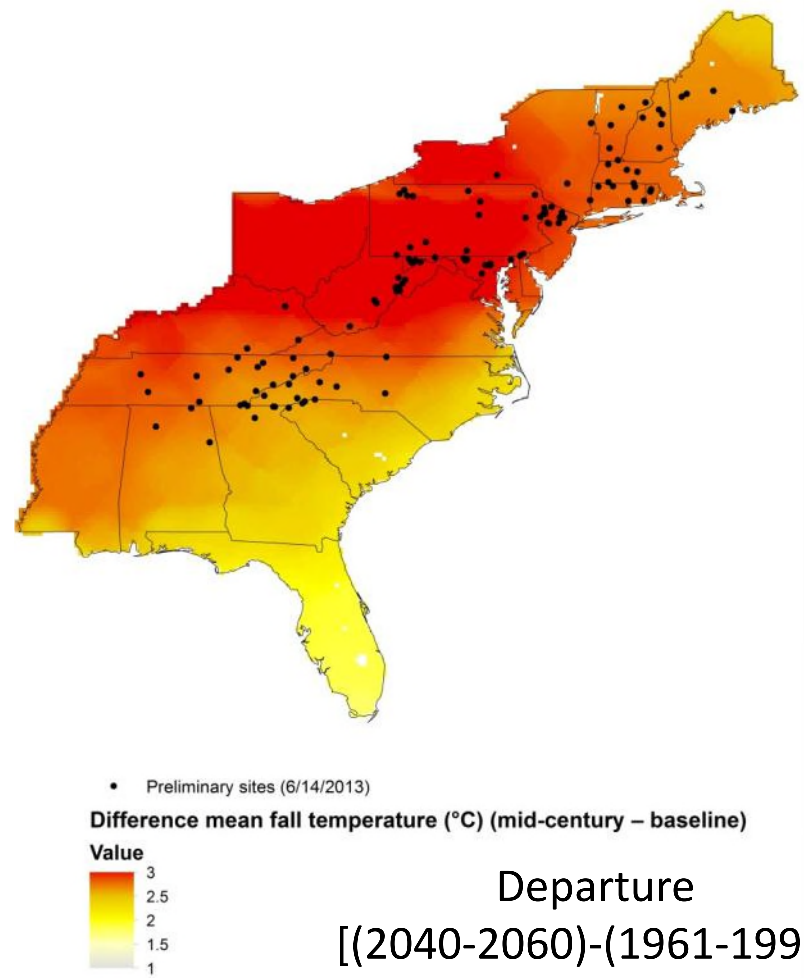
Air temperature – Mean fall (Sept-Nov)



Baseline (1961-1990)

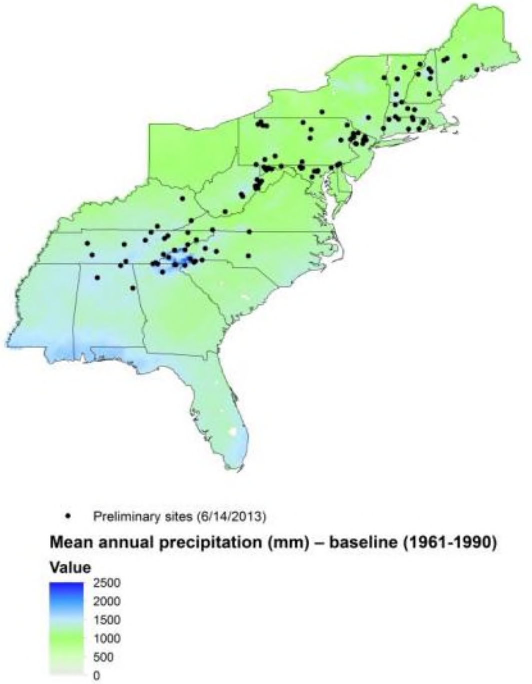


Mid-century (2040-2069)

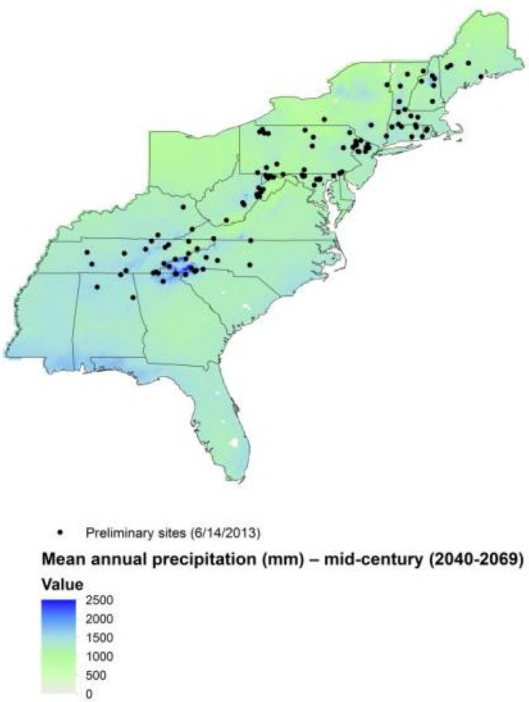


Departure [(2040-2060)-(1961-1990)]

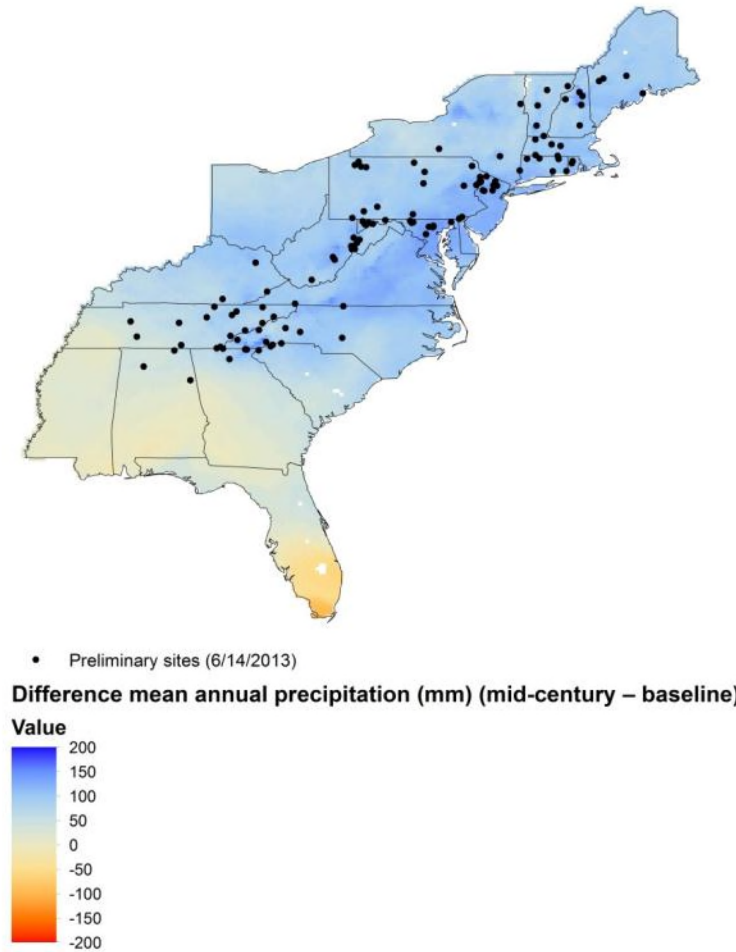
Precipitation – Mean Annual



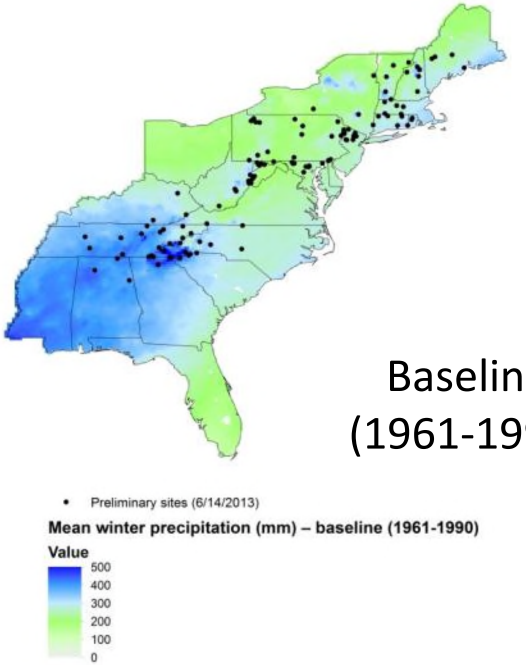
Baseline (1961-1990)



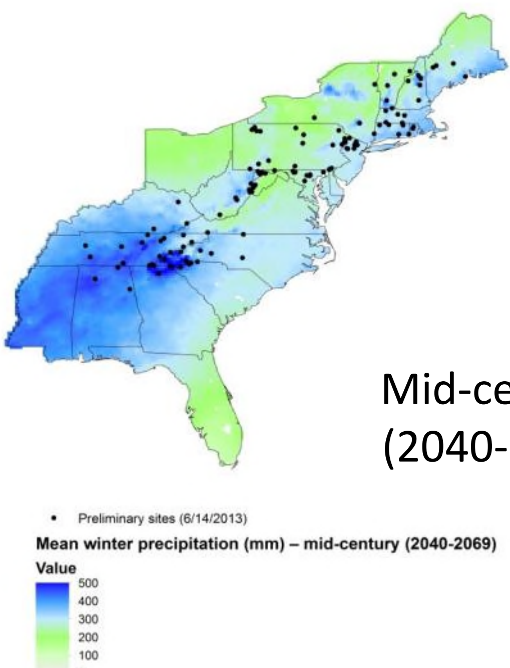
Mid-century (2040-2069)



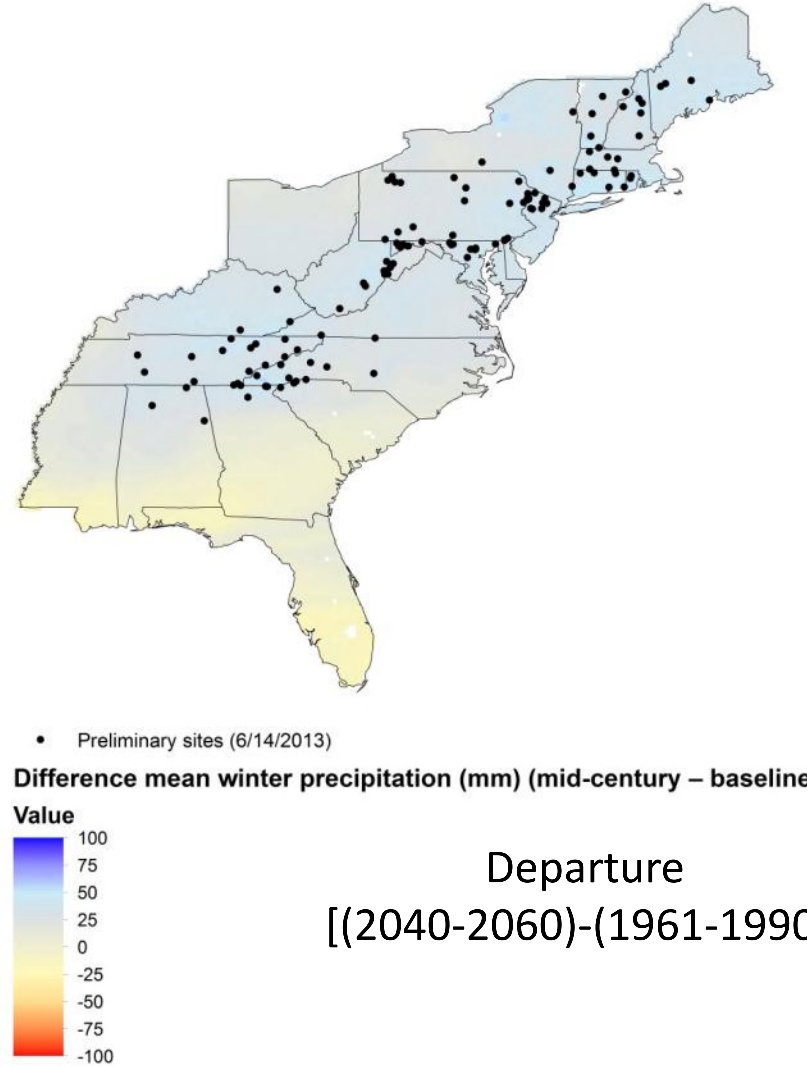
Precipitation – Mean winter (Dec-Feb)



Baseline (1961-1990)

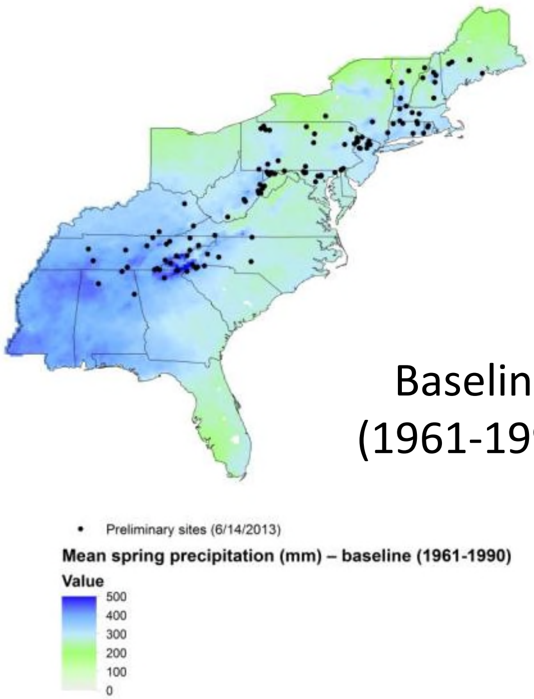


Mid-century (2040-2069)

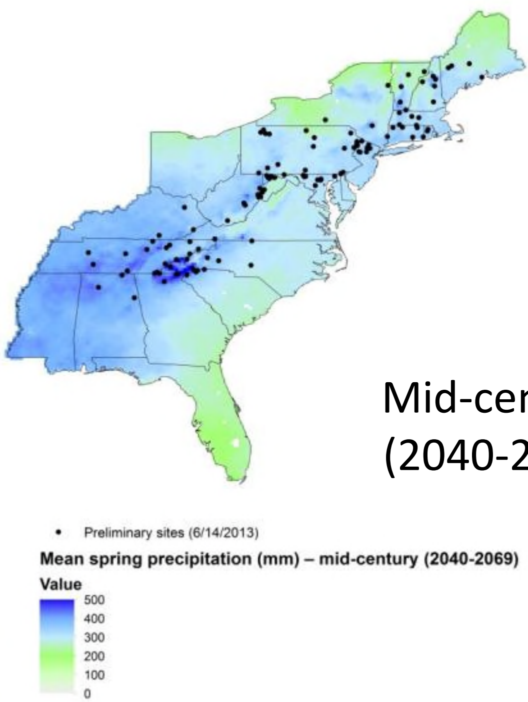


Departure [(2040-2060)-(1961-1990)]

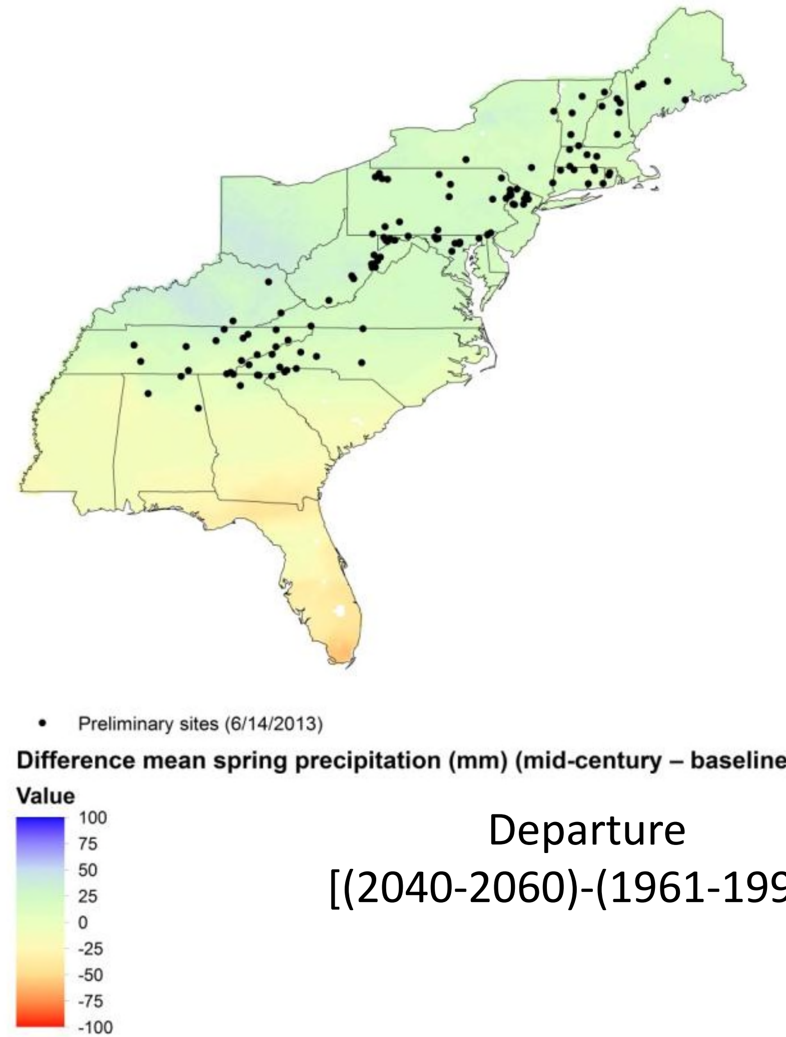
Precipitation – Mean spring (Mar-May)



Baseline
(1961-1990)

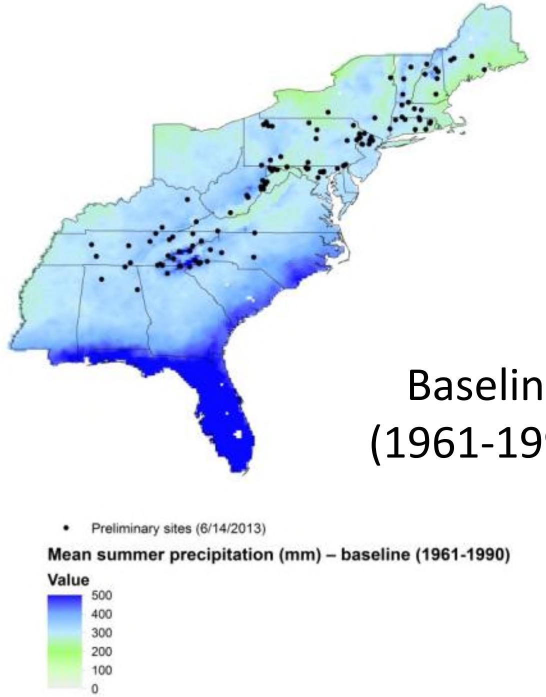


Mid-century
(2040-2069)

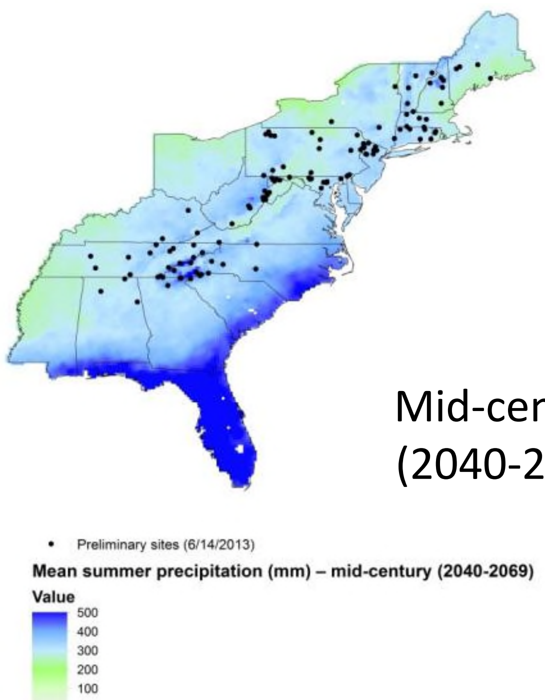


Departure
[(2040-2060)-(1961-1990)]

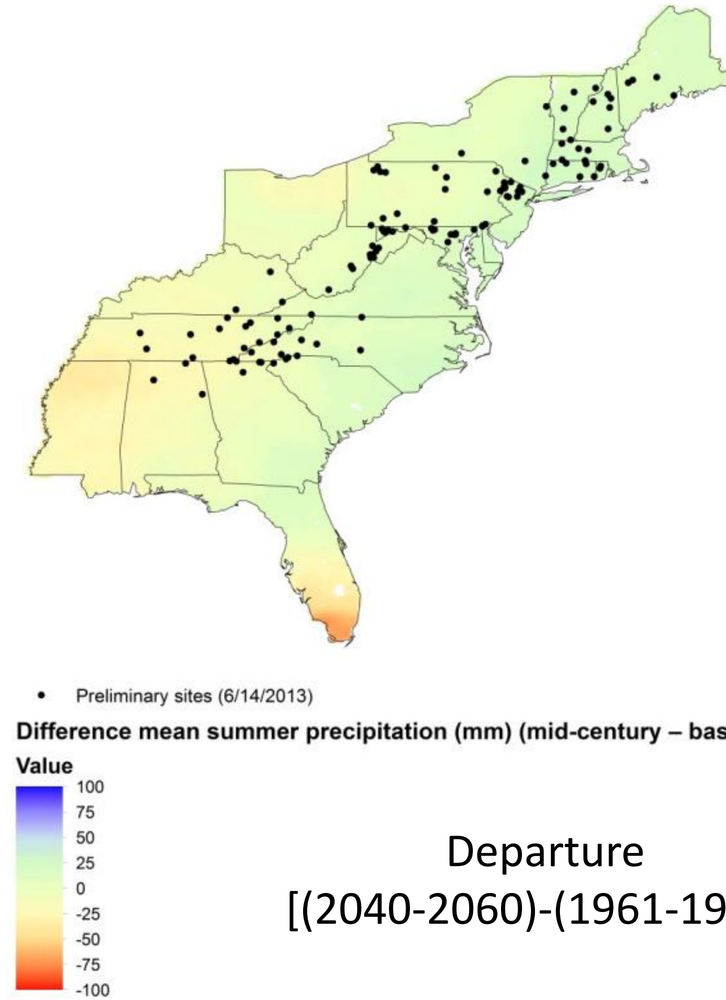
Precipitation – Mean summer (Jun-Aug)



Baseline
(1961-1990)

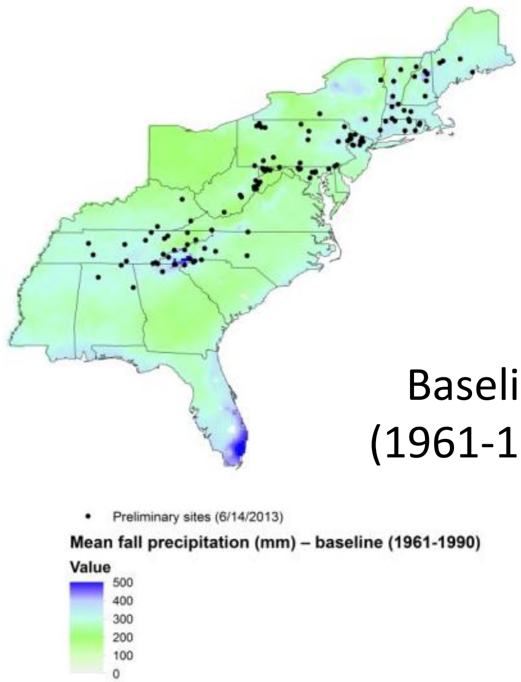


Mid-century
(2040-2069)

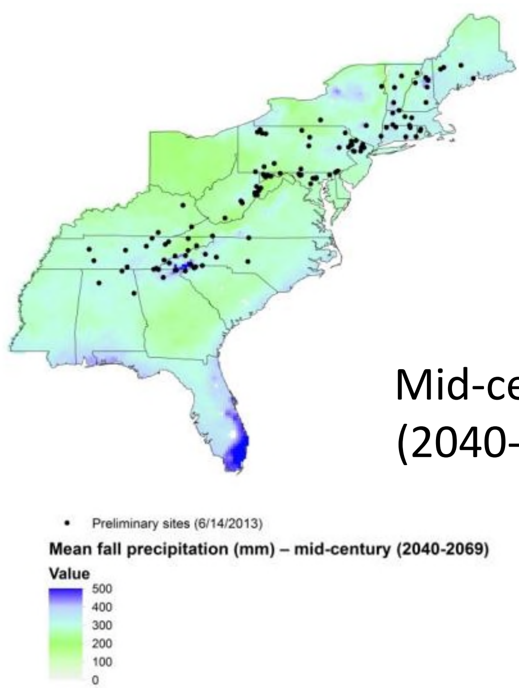


Departure
[(2040-2060)-(1961-1990)]

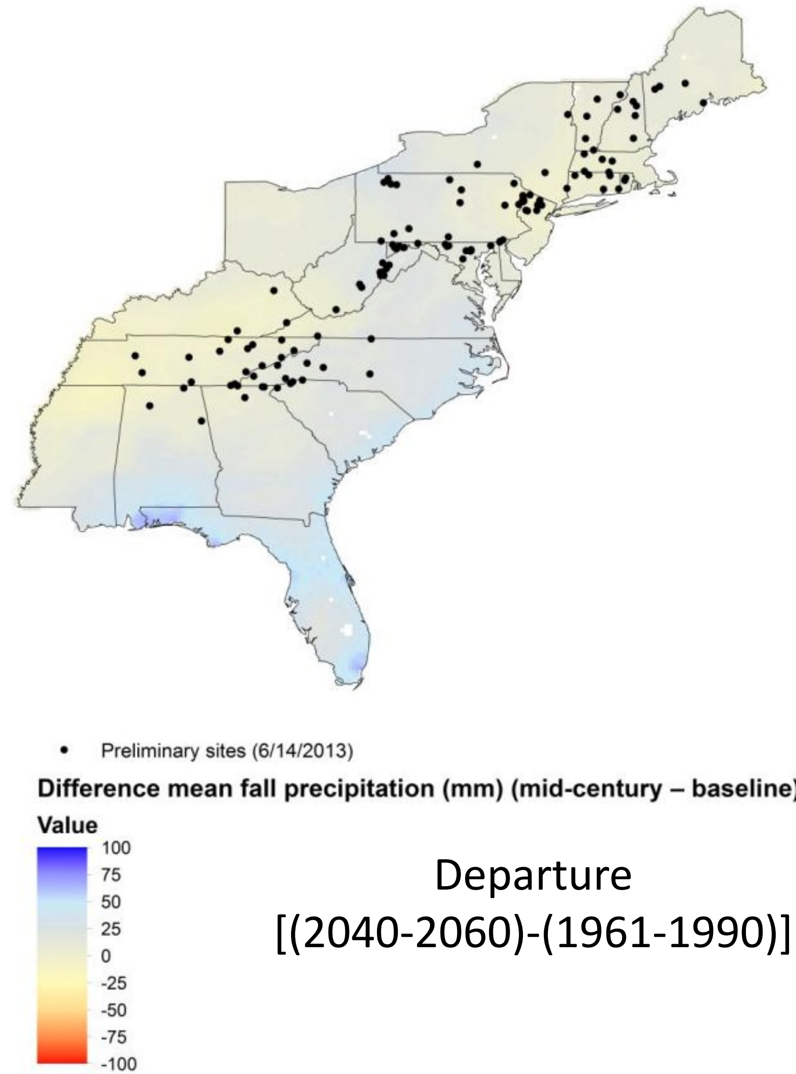
Precipitation – Mean fall (Sept-Nov)



Baseline (1961-1990)



Mid-century (2040-2069)



Departure [(2040-2060)-(1961-1990)]